International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 16 (2017) pp. 5540-5549

Discussing the Role of Data Mining in Development of Evidence Based Decision Support System for e-Healthcare

Mohammad Taha Khan[1], Dr. Shamimul Qamar[2] and Dr. Ripu Ranjan Sinha[3]

[1]Research Scholar, Suresh Gyan Vihar University, Jaipur, Rajasthan-302025, India.

Orcid: 0000-0003-3351-5433

[2]Professor, College of Computer Science, King Khalid University, Abha, Kingdom of Saudi Arabia.

[3]SS Jain Subodh PG College, Jaipur, Rajasthan-302025, India.

Abstract

A major issue highlighted in healthcare is to provide accurate and error-free diagnosis and treatment to patients. Traditionally, most of the healthcare providers individually or collectively provide their services based on their knowledge and experiences. A poor clinical decision is not acceptable in treating a patient as it can give disastrous results. To provide an error free and accurate service, healthcare providers must employ appropriate computer-based information and/or decision support systems. In present days most of the hospitals are using some sort of hospital management information systems (HMIS) to manage their healthcare services or patient data. A huge amount of data is getting generated by these systems. There is a wealth of hidden information in these data which can be very useful to support clinical decision making but unfortunately these data are not getting used for decision making process. The theme of this research is "How we can extract useful information from these data to support decision making by healthcare practitioners?". Nowadays one important approach in healthcare has been recognized as Evidence Based Medicine (EBM). EBM is use to make clinical decisions about the care of individual patients using the best available evidence. In this research the role of data mining in developing a decision support system is discussed and a case study of EBM integrated HMIS is presented.

Keywords: Data Mining; Data Warehouses; Decision Support Systems; Knowledge Based Systems; Management Information Systems; Open Source Software

INTRODUCTION

Healthcare sector is the one of fastest growing and world's second largest service sector. According to researchers and academics, the total contribution of the health sector to the gross domestic product has been estimated to be 10 percent for developed countries. The healthcare sector of developing countries like India is growing at an alarming rate and research suggests that its total worth in the year 2020 will be US $280 billion industry [Cognizant 20-20 Insights]. Information Technology (IT) has positively impacted healthcare.

The healthcare industry is considered to be complex and highly driven by globalization, intense competition and market trends along with governmental legislations and policies. Health Information System can help in comprehensive management of medical information and its dissemination among different stakeholders such as government, hospitals, patients etc. Over the time information and communication technologies (ICT) has made healthcare more accessible, interactive, and highly useful through components like telemedicine, Picture Achieving and Communication System (PACS), and Healthcare Information Systems (HIS) [23].

E-Health is based on information technology and electronic communication, with the aim of revolutionizing the healthcare organizations and the way healthcare operations work. One of the major factors that are responsible for influencing healthcare and its quality is the ability of the organization to provide accurate, previse and error free diagnosis of the disease. The cases undertaken by healthcare providers on individual or collective level vary from one another on basis of knowledge, experience and conditions. Hospitals strive for innovation and providing superior quality care to patients with the intention of improving service delivery. In the healthcare sector, high quality service indicates that patient are diagnosed accurately and properly and provided efficient treatment, which would improve their health outcomes. If the clinician had failed to diagnose the disease correctly, it will have negative outcomes and can have a negative impact on the patient's health. Thus, errors in healthcare sectors are

unacceptable. Hospitals also aim at reducing the costs of medical tests. To achieve this, computer based information systems and decision support systems are needed to be employed. Consequently, hospital management information system can help in managing patient data efficiently. However, it should be noted that patient data is very rarely used during the decision making process of the clinician. This type of data can be effectively used to avoid medical errors and reduce operational costs significantly. Healthcare organizations generate and collect large amount of data. This situation raises an important question: "How useful information can be extracted from this data? And how this useful information can support decision making by healthcare practitioners?" This research tries to search the answers for these questions.

Data mining is an evolving area in information technology because of the ready availability of large quantities of data. An era of open information is now under way. According to IBM, 100GB data was getting generated a day globally in 1992, now in 2017 the total amount of data that is generated on per second basis has been estimated to be 50000GB. Furthermore, it is suggested that the 90% of the data that is found today, had been generated in the last decade because of innovation in information communication technology [8]. Datasets have been created that cover all areas of human endeavor, including business, medical and clinical, geographical, and image data. Data mining has been identified as the technique that is used to take out information, which is hidden from the large pools of database (Han et al., 2012). It aims to automatically extract knowledge in an explicit form from large scale data. This suggests that that the extracted information and data must be useful to achieve maximum efficiency.

Extracted knowledge can be proved very useful in decision making and improving the healthcare.

Nowadays one of the most important approaches recognized in healthcare for decision making is Evidence Based Medicine (EBM).EBM is to use the best available evidence in making decision for care of individual patient. In form of EBM modern healthcare has got a new direction.

The task of EBM is defined as to avert, identify, analyze and propose medical treatment of diseases in an efficient manner by using medical records and data available in the form of medical evidence [27]. External and internal clinical knowledge is needed to provide the medical evidence efficiently and timely manner. External evidence based knowledge fails if the patient's health condition is not adjusted. Consequently, internal evidence is needed to fill this gap and to avoid medical errors. As per studies conducted by the Institute of Medicine in America it is suggested that every year, 100,000 Americans lose their lives because of medical errors. The same studies also suggest that 80% of the medical errors occur because of wrong data entry.

This paper investigates role of data mining in decision support in healthcare and will discuss a case study of one advance pathology management system (APMIS) which is integrated with some internal as well as external sources of evidences for decision support at point of reporting.

TRANSFORMING HEALTHCARE THROUGH DATA WAREHOUSING AND DATA MINING

Data warehouse and data mining are recently gaining popularity in information technology rapidly. As use of IT is growing day by day in healthcare industries, the challenges are rising for managing data in proper way so it can be used in future. The data warehouse is considered to be the central area, where data is stored from several systems and needed for decision-making across the whole organization. Integrating these heterogeneous data sources into a consistent framework is a major challenge. Data mining is defined as the process that is used to extract hidden information from given data set. It is considered to be an important approach, which is used by businesses to aid their decision making process.

The goal of this section is to discuss the concept of data warehouse and data mining.

Some people are confused with data warehouse with online transaction processing (OLTP) but these are two different terms with a little bit common as both can be used for getting the useful information from data. First of all W. H. Inmon [1992]: coined the term data warehouse "A warehouse is defined as a subject oriented, integrated, time-variant, and non-volatile collection of data which can be used for supporting of decision making management process". As compared to the conventional On-Line Transaction Processing (OLTP) applications, decision support has distinct requirements in terms of database technology. The goal of data warehouses to is to store data, improve functionality and ability to respond to queries. In the case of OLTP applications, these functions are inadequate. Data warehouses also store data, which is analyzed and consolidated over given period of time. The size of data warehouses can be huge varying from and is expressed in gigabytes to terabytes. It is important that the data is available at the right time for decision makers to aid their decision making. Complex queries are also submitted in the data warehouse. These queries review several records. Simultaneously, the submitted queries can connect, examine and assemble operations in the warehouse. The main issues in the multi-user decision support system identified by literature are the time of responding the query and the throughput of the query.

Fig. 1 depicts the architecture of the DWH. It demonstrates the processes that are involved. It starts with data collection and ends with the delivery of the data to end user: the decision makers. The data source is saved in different formats, depending on the source system. The data collection is

achieved either by external system or operational systems.

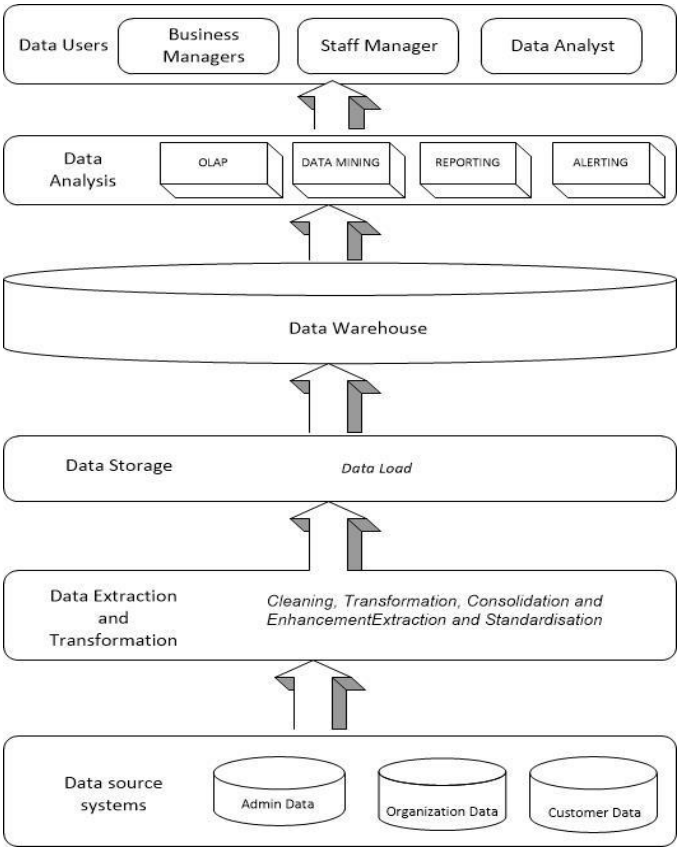


Figure 1: Data Warehouse Architecture.

A. Use of Data Mining Techniques in healthcare

This section aims to define in what way various data mining techniques are utilized with respect to the medical data in order to evaluate the required information. As we have already discussed that as use of ICT is rising in healthcare, day-by-day and same time medical databases are growing [1-2]. This growing network in the medical databases is quite encouraging for the medical researchers in order to make them use data mining for the respective databases to obtain knowledge. In case where the volume is increased in regards to the stored data, these techniques of data mining play a significant role in order to find the related patterns as well as mining the knowledge so that the better patient care can be provided with effectual diagnostic abilities. With respect to the data mining techniques, we observed the several serious questions' solutions, such as:

1. What action is required to be undertaken in order to improve the patient treatment for the records given for dialysis patients.
2. Which treatment is more useful to help cancer patient grounding on the historical patient records provided, chemotherapy separately, radiation separately, or in case of both radiation and chemotherapy?

Customary statistical methods are to be used in order to calculate the data patterns in case the data set is small; such conventional methods are not applicable in proper working with respect to huge data set. Thus, the data mining is helpful to extract the hidden patterns. It facilitates with the automatic pattern identification and more importantly endeavors to search the patterns from such large amount data related to which it is hard to find in collaboration to the statistical methods.

Data mining entails the techniques set not a single. It encounters the idea behind what is presenting as well as finding more knowledge that is hidden.

Data mining groups the variegated techniques and approaches, which are used for various purposes in knowledge discovery. Statistical techniques, machine learning, visualization, etc., are the base for the data mining techniques and methods. Data mining models are classified into two categories, which are discussed as follows:

Descriptive Models are used to identify the data patterns with the help of rules of association, visualization, clustering techniques and pattern recognition.

Predictive Models are used for predicting the data. Classification and regression analysis of the data along with time series analysis is used for prediction.

Classification: It is an important task of predictive modeling of data mining. This is use to predict the class of a data, it classifies a data item into one of different classes, these classes are predefined. The classification model is based on classification rules. The training data is used to understand the data. The rules of classification also focus on understanding the data of each class, which is found in the database.

Let's have an example from healthcare, from known cases of disease classification rules can be extracted. It can be used to aid diagnosis of diseases, based on the symptoms that have been identified.

One of the most important applications of classification is medical diagnosis. A patient with heart problems may be classified based on various types of heart diseases.

For applying the classification technique prior knowledge of under consideration data is necessary [2].

Suppose we have database of patients D. We may regard this patient's database (D) as set of sequences (x1, x2 … xn), x1, x2 … xn being values of attributes A1, A2 … .An of a particular disease. Different classes C= {C1, C2, C3 … Cn} of patients can be established based on gravity or specific classification type of the disease.

Basically, the problem concerns definition of a function f=D→C such that each ti€ D is mapped to f(ti) corresponding to some Cj.

Regression: Another task of predictive modeling of data mining is regression. It is a method which use some known type of function to map target data. It is used to determine the output value on basis of the input values.

Time series analysis: Time series analysis focuses on analyzing data for a given period of time, which is spaced. For instance, the characteristics or features of the diseases can be undertaken on hourly or daily basis. However, this depends on the existing condition of the patient. This analysis is applied for the purpose of predicting future values or for establishing similarity between time intervals.

Predictive Modeling: As we have already discussed in above section that classification, regression, time series analyses are major tasks of predictive modeling. On basis of past and present data values, it determines that provides forthcoming data trends. Predictions may be made using time series analysis, regression analysis, or other statistical techniques.

Visualization Technique: It is considered to be a descriptive model of data mining, which focuses on identifying data patterns for data sets given. Interesting subsets of medical data sets can be identified by using scatter diagrams in a Cartesian plane of two medical traits of interest. For instance, interesting subsets related to blood sugar can be found in patients with heart diseases.

Association Rule: This technique focuses on identifying the association between two quantities. It is expressed as "A1 ^A2 …^ Ai => B1 ^ B2^…^ Bj", that is, B1, B2, B3, …and Bj turn up with objects A1, A2, A3 … and Ai in the target data.

ROLE OF DATA MINING AND WAREHOUSING IN FASCILITATING EVIDENCE BASED MEDICINE

From hundreds of the years testing the outcome of medical interventions has been performed. All the fields of welfare and healthcare started getting impacted by this effort during last century. The founder and the pioneer of the evidence based practice was Archie Conchrane, who was a Scotland based epidemiologists. His research focused on prompting evidence based medicine. He indicated the medical evidence is important for improving the effectiveness of treatments in randomized control trials, which is defined as the scientific inquiry to test medical treatments or medical drugs. RCT is considered to be an important source of scientific inquiry and thus, has high degree of reliability and validity [24]. His work had been popular in promoting EBM and thus, the Cochrane Centers had been established to research on the topic. The Cochrane Collaboration has also been formed, whose primary duty is to improve coordination and create promotion of EBM in different regions and communicating with the public. It [Cochrane Collaboration, 2007] is a world-wide endeavor dedicated to tracking down, evaluating, and synthesizing RCTs in all areas of medicine. It is a major force in the EBM movement. The Cochrane Collaboration provides the Cohrane Library [Cochrane Library, 2007], a collection of medical databases that contain high-quality, regularly updated independent evidence to support healthcare decision making. There are numbers of reliable evidences from Cochrane, clinical trials, other systematic reviews, and more in Cohrane Library.

David Sackett and Gordon Guyatt conducted a research at McMaster University, in Canada. They were responsible for developing the explicit methodologies, which were used to analyze and identify the "best evidence". According to this group, in the making decision for the care of individual patient, conscientious, careful and previse usage of the evidence, which is deemed to be the best evidence, is called evidence based medicine. Consequently, it is suggested that EBM uses external and internal clinical data to provide the best quality care to the patients.

Evidence-based medicine has been emerged as a new direction in e-healthcare. The major task of EBM is prevention, diagnosing and medicating diseases using medical evidence [24].

External clinical expertise and internal clinical expertise are equally important to get most reliable evidence for a given disease. Clinical expertise, both in-house and external, should be conveniently available to healthcare practitioners. It is not advisable to apply external evidence-based knowledge to the patient without tailoring it to the patient's health condition. One of the major research challenges is the integration of externally obtained evidence based data into in-house information system and using it to find suitable therapy for a particular patient having a particular disease. Basic evidence based care is shown in Fig. 2.

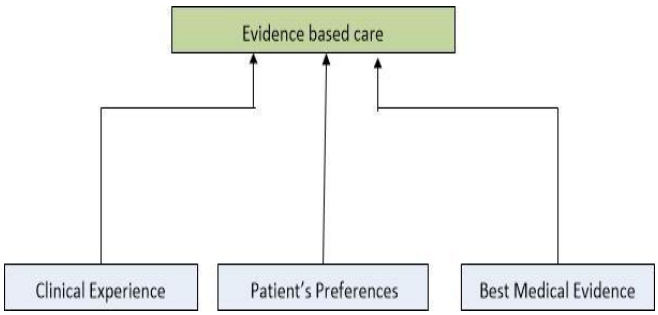


Figure 2: Evidence Based Care

A data warehouse system that makes evidence-based medicine possible is an invaluable, powerful and easy-access tool for making strategic decisions and has relevance for acceptance and practice of evidence-based medicine. The practice of evidence-based medicine is heavily dependent on IT support as gathering all the necessary knowledge about a particular disease is a difficult task for clinicians without IT support.

The importance role of decision support system is increasing with drastic changes in the field of healthcare. For strategic decision-making many of the healthcare institutions are deploying data warehouse and data mining based applications as decision-support tools.

In today's health industries, the combination of data warehousing and data mining technologies along with evidence-based medicine provides a new field for application of information and communication technologies. Reduced treatment costs and increased healing rate of patients are the primary interest of most of the healthcare entities like medical institutions, as well as health insurance companies. In the area of EBM applications based on data warehouse and data mining can help in avoiding duplicate examinations, automation of routine tasks and simplification of the accounting and administrative procedures. In the long term, these applications could be proved very economical and overall cost of treatment can be reduced up to certain level.

One of the major concerns of healthcare industry of present days is chronic illnesses caused by the growing and aging population. Diseases of elderly people, like Diabetes, Alzheimer's disease, cardiac insufficiency, and sight loss (macular degeneration) will cause more treatment effort and therapy costs than the treatment of most difficult illnesses (cancer and heart attack) generate nowadays. If these diseases are detected in the early stages, treatment of these diseases can be more efficient and more cost-effective. To recognize the patterns of disease formation and development is the mission of modern medicine. EBM's major task is to analyze the existing medical records, clinical studies and to search for the recurring samples of disease symptoms.

Data warehousing and data mining techniques are integral part of evidence based medicine. Acquisition and gathering of existing medical experience from diverse data sources, and statistical analysis of that data can be handled by data mining and data warehouse techniques.

Evidence based guidelines (EBGs) are extracted experience values and formulated knowledge. Evidence-based guidelines are used for predicting more efficiently, discovering, and providing a better treatment of diseases [27].

In this section, some application fields which are relevant to the clinical knowledge management are presented, especially:

- Developing new knowledge – the data warehouse and data mining supported creation of evidence-based guidelines and clinical pathways
- Knowledge sharing – the data warehouse as an easy to use platform for knowledge dissemination among healthcare decision makers.

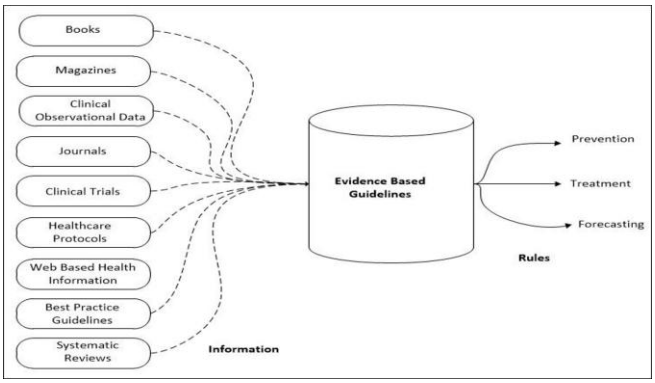


Figure 3: Data flow in EBM

Practicing EBM: The major steps in practicing evidence-based medicine are listed below [24]:

1. Formulating proper questions
2. Finding the available evidence.
3. Evaluating the evidence under real circumstances for utility and validity
4. Administering the evidence selected
5. Evaluating the performance

Evidence-based medicine is needed for:

- Making a diagnosis:

  Evidence based medicine can help a clinician in making the right diagnosis.

- Administering appropriate therapy to patients:

  EBM guidelines will offer different treatments from data-base, the one best suited to the patients, keeping in view health risks and preference of patient, can be applied.

- Making a prognosis:

  Having information of various similar cases at hand, clinician can make educated prognosis of a patient's future state of health.

- Clarifying an etiology:

  EBM provides expert knowledge about the possible causes of a particular medical condition. For example, it may help find the correlation between consumption of nicotine and heart disease.

  Fig. 4 shows EBM decision support at pint of care.

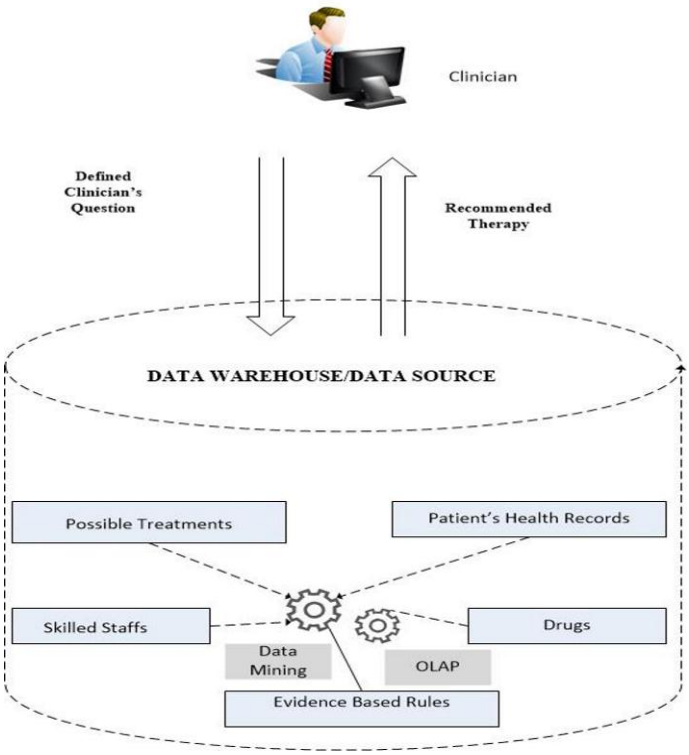


Figure 4: EBM Decision support at point of care

CASE STUDY: EBM INTEGRATED HMIS

This case study discusses the development of an Advanced Pathology Management Information System (APMIS) which uses EBM to provide support at the time of reporting. To help in decision making at the time of reporting APMIS is integrated with external evidence sources as well as some internal evidences. First section of this case study discusses how classification techniques can be used on some online available data [15] to mine rules for predicting breast cancer and second section is detail description of development of open source APMIS integrated with external evidence sources like MEDILINE and Cochrane database of systematic reviews.

A) *Rule Mining from Online Cancer Data:*

Classification is a widely applied tool of predictive modeling. This technique is used in healthcare to predict the diseases based on symptoms. Target classes are divided into data samples. Target class for each data points is then predicted [2].

In this section rules generation using data mining is discussed here. Classification algorithms C4.5 and C5.0 can be applied on health data to generate the decision tree, which is considered to be an important aspect of discovering the knowledge. The decision tree model is considered to be based on predictive modeling and thus, focuses on classifying the data in the form of tree.

Online data of breast cancer available at University of Wisconsin Hospitals, Madison (Dr. William H. Walberg) [15] is used for the breast cancer prediction case study.

i). Breast Cancer data Attributes:

Total Cases:	*599*
Attribute	*Domain*
1. Sample code number	id number
2. Clump Thickness	1 – 10
3. Uniformity of Cell Size	1 – 10
4. Uniformity of Cell Shape	1 – 10
5. Marginal Adhesion	1 – 10
6. Single Epithelial Cell Size	1 – 10
7. Bare Nuclei	1 – 10
8. Bland Chromatin	1 – 10
9. Normal Nucleoli	1 – 10
10. Mitoses	1 – 10
11. Class:	(2 for benign, 4 for malignant)

ii). Specification of Attributes:

The target attribute:	Class
Sample code number:	ignore
Clump Thickness:	continuous
Uniformity of Cell Size:	continuous
Uniformity of Cell Shape:	continuous
Marginal Adhesion:	continuous
Single Epithelial Cell Size:	continuous

Three types of files are supported by C4.5 and C5.0.Name of classes, attributes and attribute values are provided by names files. Training cases for decision tree generation are described by data file. Target attribute (a class that can only have one of the two values) has either value 2 (Benign) or value 4 (Malignant). Benign means non-cancerous, while malignant means cancerous.

Malignant tumors are capable of invading and destroying nearby tissues and spreading to other body parts, while benign tumors do not spread.

The values of 2 and 4 are given to class attributes to avoid confusion with values of other attributes. There are several other attributes (e.g the ones mentioned above) which can have values from1 to 10.

Decision Tree and Rules Generated:

A tree of size 29 with 5 train error is generated using C4.5 algorithm shown in Fig. 5.

By 5 train errors it is meant that there were five cases where error was noted down after 400 records were run on C4.5.

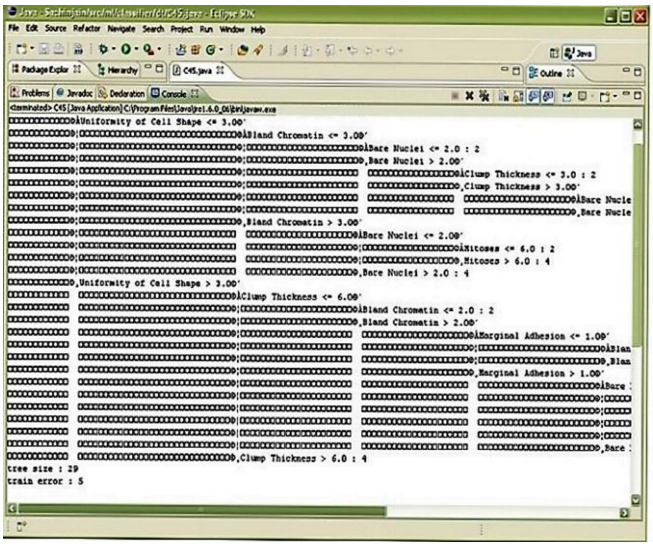
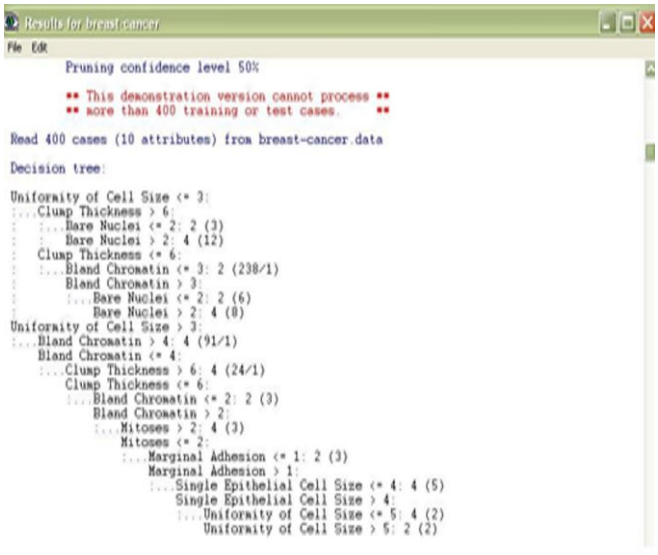


Figure 5: Tree produced before pruning using c4.5

To reduce the complexity of classifier and improve the predictive accuracy, tree size is reduced to 17 after pruning.

Fig. 6 shows the tree produced after running C5.0, reading 400 cases having 10 attributes.



```
Results for breast-cancer
File Edit

      Pruning confidence level 50%

      ** This demonstration version cannot process **
      ** more than 400 training or test cases.     **

Read 400 cases (10 attributes) from breast-cancer.data

Decision tree:

Uniformity of Cell Size <= 3:
:...Clump Thickness > 6:
:   :...Bare Nuclei <= 2: 2 (3)
:   :   Bare Nuclei > 2: 4 (12)
:   Clump Thickness <= 6:
:   :...Bland Chromatin <= 3: 2 (238/1)
:       Bland Chromatin > 3:
:       :...Bare Nuclei <= 2: 2 (6)
:           Bare Nuclei > 2: 4 (8)
Uniformity of Cell Size > 3:
:...Bland Chromatin > 4: 4 (91/1)
    Bland Chromatin <= 4:
    :...Clump Thickness > 6: 4 (24/1)
        Clump Thickness <= 6:
        :...Bland Chromatin <= 2: 2 (3)
            Bland Chromatin > 2:
            :...Mitoses > 2: 4 (3)
                Mitoses <= 2:
                :...Marginal Adhesion <= 1: 2 (3)
                    Marginal Adhesion > 1:
                    :...Single Epithelial Cell Size <= 4: 4 (5)
                        Single Epithelial Cell Size > 4:
                        :...Uniformity of Cell Size <= 5: 4 (2)
                            Uniformity of Cell Size > 5: 2 (2)
```

Figure 6: Rules Generated using c5.0

Case study shows C5.0 handles missing values easily but C4.5 shows some errors due to missing values.

B) Design and Development of Open Source APMIS Integrated with External Source of Evidences

This open source low cost APMIS development is based on CARE2X customization. AMPIS is deployed at one of the leading client hospital in New Delhi, India. Statistics of client hospital are given in Table 1.

Table 1: Details of Client Hospital

List of Provisions	Totals (in numbers)
Beds	110
Medical departments	5
Diagnostic laboratories	6
Surgical departments	2
Operating rooms	4
Nursing wards	16
Central Pharmacy	1
Central material departments	1

Histopathology and Cytopathology are using APMIS for diagnosis purpose. APMIS covers the complete workflow of Lab and provide advance features like automatic accession key generation, sample tracking, interesting report management, disposal management, advance record search, role base authorization for security purpose and decision support at the table of reporting. Multiple users like doctors, technicians and data entry operators are using the APMIS simultaneously at different level.

CARE2X: This system is based on the open source technology LAMP [12]. The system is devised as a solution to fulfill the objectives of centralized database server, business logic and multi-location clients who can access the web based application through Intranet. The software architecture is a versatile, request/response based and object-oriented infrastructure that is intended to improve usability, flexibility, interoperability and scalability.

CARE2X is one of the most popular open source healthcare systems. CARE2X Integrated Healthcare Environment (IHE) and it is fully web based system. As it an integrated healthcare environment it integrates the data, information, functions and workflows in one single environment.

It can also handle non-medical services (i.e. maintenance) easily. CARE2X is the modular, highly scalable and web based solution. CARE2X uses server side scripting language for its development so no changes on web browser are needed for module updates or extensions, which means no problem of

network interruptions and downtimes. Multiple server configurations are supported by CARE2X design to for efficient distribution of traffic. It is released under General public license GPL v2.0.

During the last few years the development efforts of the original free version have been concentrated in developing countries [9], which are also seen in the download statistics. Out of 16500 downloads to 150 countries in the last year 85% goes to developing countries.

That means acceptability of open source software in developing countries is rising exponentially. It is very easy to download CARE2X and use it.CARE2X has a great potential to fulfill the requirement of hospitals in developing countries where cost is more important.

It almost consisting every features which is required in a complete healthcare solution.

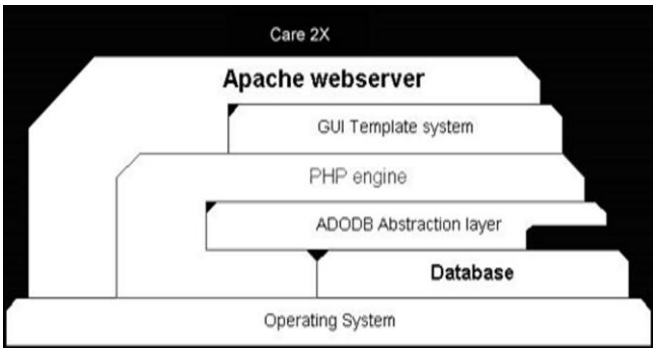


Figure 7: CARE2X architecture [12]

As MVC architecture is used for the system development, it is easy to maintain quality and business requirements for the user. HTML, JavaScript and CSS have been used to implement presentation tier, PHP to implement middle tier and MySQL to implement Database layer. Architecture of CARE2X is given in Fig. 7.

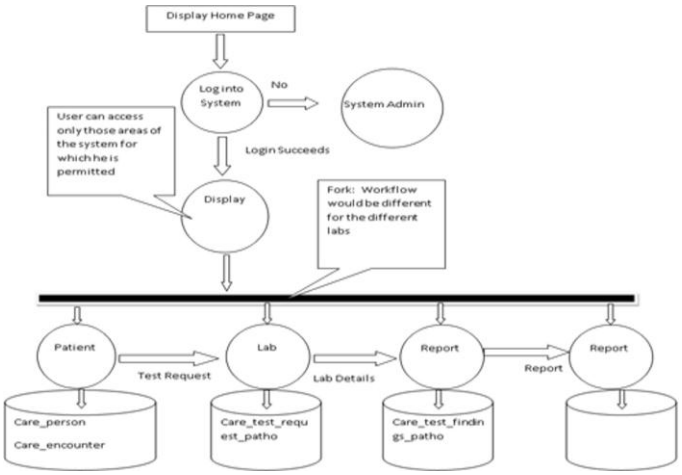


Figure 8: APMS DFD

C) Integration of External Evidence Sources in APMIS

There are a lot of research work has been done for treatment of different diseases, number of proved evidences are available online.

MEDLINE (Medical Literature Analysis and Retrieval System) is a bibliographic database of biomedical and life sciences, which include citations for articles from scientific journals of medicine, health care, biochemistry, biology, and microbiological evolution, accessible through the PubMed search interface from the NLM [28]. PubMed includes lists of books as well as the journal articles of MEDLINE. Some rules generated in case studies (above discussed) as well as Cochrane library [29] and MEDILINE, the sources of external evidence based guidelines have been integrated with APMIS to support the decision making and conflict management at point of reporting/care.

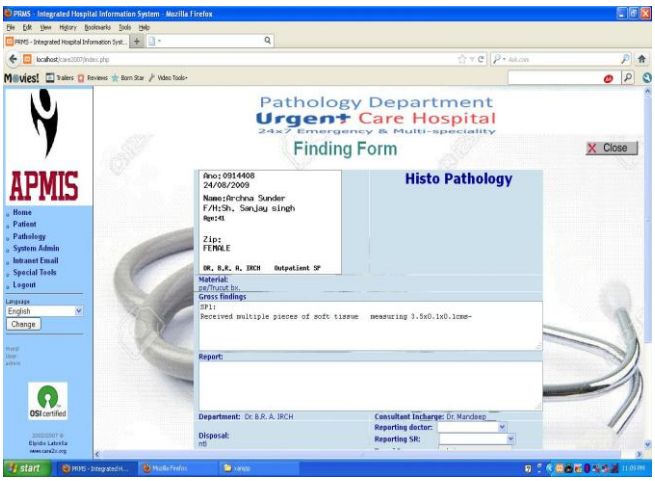


Figure 9: Reporting in APMIS

Fig.9 show reporting form in APMIS and Figure 10 shows the use of MEDLINE for support.

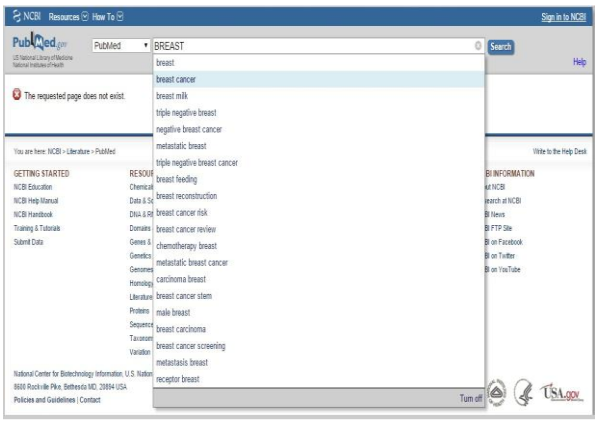


Figure 10: Using MEDLINE as external evidence

CONCLUSION

It has been shown that data mining and data warehousing technologies can play a transformative role in healthcare sector. These technologies can play equally important role in development of an endurable decision-support system in the field of healthcare. The requisites for the EBM integration into decision making process have also been discussed. Application of data mining and data warehousing (DWH) technology can accelerate creation of rules pertaining to evidence-based therapy and ensure that recommendations made are correct.

EBM is a new direction in e-healthcare. EBM is just about justified use of proved external evidence with help of internal expertise at right time and right manner for improving the patient care. Major advantages of applying EBM in the decision-making process are as follows:

- Reduction in conflict and medical errors
- Efficiency Improvement
- Improvement of patient care
- Treatment costs Reduction

It is also discussed that open source software is very useful as it can provide a low cost alternative of proprietary software.

By case study it has been shown that CAREX customization provides a low cost healthcare solution for a client hospital in India.

REFERENCES

[1] Ionuț ȚĂRANU, Data mining in healthcare: decision making and precision, Database Systems Journal vol. VI, no. 4/2015.

[2] Joshi S., Nair M.K. (2015) Prediction of Heart Disease Using Classification Based Data Mining Techniques. In: Jain L., Behera H., Mandal J., Mohapatra D. (eds) Computational Intelligence in Data Mining - Volume 2. Smart Innovation, Systems and Technologies, vol 32. Springer, New Delhi.

[3] Pravej Ahmed,Saqib Qamar,Syed Qasim Afser Rizvi: Techniques of Data Mining In Healthcare: A Review, International Journal of Computer Applications (0975 – 8887) Volume 120 – No.15, June 2015.

[4] Mahaveer Golechha,'Healthcare agenda for the Indian government', Indian J Med Res 141, February 2015, pp 151-153.

[5] Ted Epperly, Richard Roberts, Salman Rawaf, Chris Van Weel, Robert Phillips, Juan E. Mezzich, Yongyuth Pongsupap, Tesfamicael Ghebrehiwet, James Appleyard " Person- Centered Primary Health Care:Now More Than Ever", The International Journal of Person Centered Medicine, Vol 5, No 2 (2015), ISSN: 2043 7749.

[6] M. Glen Doss, Healthcare Turns to Open Source Software, eHealth care strategy and trends, Volume 16 Number 5 - May 2014.

[7] Paul C. Zikopoulos, Chris Eaton, Dirk deRoos, Undestanding Big Data-Analytics for Enterprize Class, Mcgraw Hills.

[8] Sherif Sakr, Big Data 2.0 Processing Systems: A Survey, Springer Briefs in Computer Science.

Alireza Amrollahi, Mohammad Khansari, Amir Manian, Success of Open Source in Developing Countries: The Case of Iran, International Journal of Open Source Software and Processes, 5(1), 50-65, January-March 2014.

Patka, S., et al.: Recent trends and rapid development of applications in data mining. IOSR J. Comput. Sci. (IOSR-JCE) 73–78. e-ISSN: 2278-0661, p-ISSN: 2278-8727 (2014). World Health Report, https://en.wikipedia.org/ wikiWorld_Health_Report. (Accessed 30 March 2017)/

[9] CARE2X, an open source project, http://www.care2x.org/ (Accessed 30 March 2017).

[10] PHP an open source widely used web development language, http://www.php.org/(Accessed 31st March 2017).

[11] MySql Largest Open Source Database used by many renowned leading organizations, http://www.mysql.com/ (Accessed 31st March 2017).

[12] Breast Cancer Dataset: http://archive.ics.uci.edu/ml/ datasets/Breast+Cancer+Wisconsin+(Prognostic) (Accessed 31st March 2017).

[13] Tomar, D., Agarwal, S.: A survey of data mining approaches for healthcare, Int. J. Bio-Science and Bio-Technology 5(5), 241–256 (2013).

[14] El-Sappagh, S.H., et al.: Data mining and knowledge discovery: applications, techniques, challenges and process models in healthcare. Int. J. Eng. Res. Appl. (IJERA) 3(3), 900–906. ISSN: 2248-9622 www.ijera.com (2013).

[15] Simon M Karume, Samuel Mbugua ,Trends in Adoption of Open Source Software in Africa, Journal of Emerging Trends in Computing and Information Sciences, VOL. 3, NO.11 Nov, 2012 ISSN: 2079-8407.

[16] H. Durrani, S. Nigar, S. Tirmizi, S. Khoja , Collaborative Efforts towards eHealth Adoption in Asia ,Global Telemedicine and eHealth Updates: Knowledge Resources Vol. 5, 2012.

[17] A-B. M. Salem, Intelligent Technologies and Methodologies for Medical Knowledge Engineering,

Global Telemedicine and eHealth Updates: Knowledge Resources Vol. 5, 2012.

[18] Jiawei Han,Micheline Kamber,Jian Pei (2012),Data Mining Concepts and Techniques,3rd ed,Morgan Kaufmann,USA.

[19] Hauge, Ø., Ayala, C., & Conradi, R. (2010). Adoption of open source software in software-intensive organizations - A systematic literature review. Information and Software Technology, 52(11), 1133-1154. doi: DOI: 10.1016/j.infsof.2010.05.008.

[20] Dr. S M Aqil Burney, Nadeem Mahmood, Zain Abbas, Information and Communication Technology in Healthcare Management Systems: Prospects for Developing Countries, International Journal of Computer Applications (0975 – 8887) Volume 4 – No.2, July 2010.

[21] R. B. Haynes, P. J. Devereaux, and G. H. Guyatt. Clinical expertise in the era of evidence-based medicine and patient choice. Evidence Based Medicine, 7:36–38, 2002. doi: 10.1136/ebm.7.2.36.

[22] Priya Nambisan, EMR adoption among office based physicians and practices: Impact of peer to-peer interactions, peer support and online forums, 2014 47th Hawaii International Conference on System Science.

[23] Gerald Bortis, Experiences with Mirth: An Open Source Health Care Integration Engine, ICSE'08, May 10-18, 2008, Leipzig, Germany.

[24] Bonnie Spring, Evidence-Based Practice in Clinical Psychology: What It Is, Why It Matters; What You Need to Know, Journal Of Clinical Psychology, Vol. 63(7), 611–631 (2007).

[25] PubMed search interface from the NLM (www.ncbi.nlm.nih.gov/pubmed/) (Accessed 1st March 2017).

Cochrane, Trusted evidence, informed decisions, Better health http://www.cochrane.org/ (Accessed 1st March 2017).